DTIC FILE COPY

Navy Personnel Research and Development Center

San Diego, CA 92152-6800 TN 89-7 October 1988

AD-A203 763

Brain Activity During Tactical Decision-making: V. A Cross-study Validation of Evoked Potentials as Indices of Workload

NRPDC TN 89-7

October 1988

BRAIN ACTIVITY DURING TACTICAL DECISION-MAKING: V. A CROSS-STUDY VALIDATION OF EVOKED POTENTIALS AS INDICES OF COGNITIVE WORKLOAD

Mark H. Blankenship

Leonard J. Trejo

Gregory W. Lewis

Approved and released by
J.C. McLachlan
Director, Training Systems

Approved for public release; distribution is unlimited.

Navy Personnel Research and Development Center
San Diego, California 92152-6800

FOREWORD

This report is the fifth in a series of reports examining the feasibility of using neuroelectric signals to predict decision-making of combat system operators under varying workloads. The first report (HFOSL Tech. Note 71-86-6) identified assumptions underlying this approach to the study of decision-making. The second report (NPRDC TN 88-12) provided detailed analyses of the physiological changes in brain activity that occur to an irrelevant visual probe as cognitive workload increases in a combat system simulation. The third report (in press) describes relationships between physiological brain activity and combat system simulation and on-job performance variates. The fourth report (in press) provides analyses of the physiological changes in brain activity that occur in a dual-task paradigm as cognitive workload increases in a combat systems simulation.

This report describes a cross-validation from two studies of the physiological changes in brain activity that occur as workload increases in a combat systems simulation.

Research described in this report is being performed under the work units 521-804-042.03.2 (Future Technologies-Biopsychometrics) sponsored by the Office of Naval Technology and 44-521-080-203 (Biopsychometric Assessment) sponsored by Headquarters, Marine Corps (MA).

J.C. McLACHLAN
Director, Training Systems

SUMMARY

Problem

The demands of modern combat systems have the potential for exceeding the capacity of the human to accurately process information, especially during times of great stress. The capacity of the human to perceive, integrate, remember, and use information may be challenged when the individual is flying aircraft, monitoring radar and sonar displays, or operating electronic warfare systems. Exceeding the capacity of the human operator in such situations may impair decision-making and could result in costly tactical errors.

Although much is being done to improve the reliability of combat systems, not enough is being done to improve the system operators. For these reasons, the most unpredictable element in combat systems is the human operator. Years of personnel testing have not eliminated this unpredictability. In part, this is because traditional testing methods tend to measure what a person *knows rather than how a person thinks and processes information.*

This research is driven by the Navy's need for better methods of assessing combat system operators, particularly for predicting the ability of operators to continue to make accurate decisions under heavy workloads.

Objective

This report, the fifth in a series of reports concerned with the use of neuroelectric signals to predict the decision-making performance of combat system operators, provides detailed analyses of a cross-validation from two studies which used neuroelectric signals as predictors. It is important that the neuroelectric signals be reliable measures of operator performance if they are to be used in assessing operator workload.

Approach

We conducted two studies using an anti-air combat simulation (AIRDEF) on two different populations of U.S. Marine Corps personnel. The first study (Trejo, Lewis, & Blankenship, 1987) found that certain event-related potential (ERP) amplitude measures decreased as workload increased in AIRDEF. The basic design of the experiment consisted of a single-task condition in which subjects performed AIRDEF under different levels of difficulty. ERPs were recorded to an irrelevant visual probe that was presented on the same monitor as the simulation. The second study (Blankenship, Trejo, & Lewis, 1988) also manipulated the level of difficulty of the AIRDEF task. Evoked potentials were recorded from subjects who attended to visual and auditory *oddball* stimuli (stimuli that differ on some physical characteristic, i.e., intensity, frequency, and are presented with different but complementary probabilities) but who were told to ignore irrelevant visual and auditory probe stimuli. This study employed a dual-task paradigm. As in the first study, it was found that certain ERP amplitude measures decreased as workload increased on the AIRDEF task. Both studies employed an irrelevant visual probe stimulus for ERP generation. Thus, for these irrelevant visual stimuli and associated ERPs, we conducted a cross-validation of the results.

Results

The results cross-validated at two levels of specificity. First, overall analyses of variance indicated similar findings for the factors of workload, electrode site, time, and their three-way interaction. Second, the results cross-validated in four electrode site window latency combinations.

Implications

The cross-validation supports the idea that ERPs are reliable across individuals who performed the combat systems simulation. The results also suggest that this reliability may be robust enough to be used under several different testing conditions.

CONTENTS

LIST OF TABLES

INTRODUCTION

The Navy depends heavily on written tests to evaluate personnel. The results of such tests can predict academic performance reasonably well ~~(Fishman, 1958),~~ but are less effective in predicting on-job performance (Ghiselli, 1966, 1973). There is a need for new ~~kinds of tests~~ that will supplement the ~~information derived from~~ written tests and provide an improved understanding and a more complete assessment of the unique capability of each individual (Lewis & Sorenson, 1987). In an attempt to better understand the human as an integrator and decision maker in operating systems, event-related potentials (ERPs) have been employed to assess individual brain processes and their relationship to differences in on-job performance and decision making (Lewis, 1983, 1984; Lewis & Rimland, 1979; Lewis & Sorenson, 1987).

The purpose of this report is to demonstrate a cross-study validation of the use of ERPs as indices of cognitive workload. The first study ~~(Trejo, Lewis, & Blankenship, 1987)~~ found that certain ERP amplitude measures decreased by about 40 percent as workload increased from baseline to active participation in an anti-air warfare situation (AIRDEF). The study was conducted in the summer of 1985 with a group of 30 U. S. Marines stationed at the Naval Air Station, North Island, San Diego. The basic design of the experiment consisted of a single-task condition in which subjects performed an anti-air warfare simulation (AIRDEF) task under three progressively higher levels of difficulty. ERPs were recorded to a visual stimulus that was presented on the same monitor as the simulation, but this stimulus had no task relevance. It was an "irrelevant" probe which subjects were told to ignore.

The second study also used the AIRDEF task and was conducted with a separate group of 65 U. S. Marines stationed at Camp Pendleton in the summer of 1986. This study also manipulated the level of difficulty of the AIRDEF task. However, in that study, the order of difficulty levels was counterbalanced across subjects. This study employed a dual-task paradigm. Evoked potentials were recorded from subjects who attended to visual and auditory oddball stimuli but who were told to ignore irrelevant visual and auditory probe stimuli. As in the first study, it was found that certain ERP amplitude measures decreased as workload increased on the AIRDEF task (Blankenship, Trejo, & Lewis, in press).

Both studies employed an irrelevant visual probe stimulus for ERP generation. Thus, for these stimuli and associated ERPs, we proceeded to compare the results of the studies for purposes of cross-validation.

METHODS

Complete descriptions of the methods used in the two studies to be compared have been published (Trejo et al., 1987; Blankenship et al., in press). In this section we briefly cover the essential methodological details and point out major methodological differences between the two studies.

Subjects

The first sample consisted of 30 U. S. Marine Corps security personnel at the Naval Air Station at North Island (NASNI). The second consisted of 65 non-commissioned officers attending Leadership School at Camp Pendleton (NCO-CP). All subjects signed a consent form and filled out a brief biographical questionnaire. The questionnaire results from both studies indicated that the subjects were mostly right-handed Caucasians in their early twenties. Most subjects reported that they were not on medication, were not tired or drowsy, and had little sight or hearing difficulty. About half of the subjects smoked or chewed tobacco.

Air Defense Simulation (AIRDEF)

AIRDEF simulates anti-air warfare as fought from a ship (after Kelly, Greitzer, & Hershman, 1981; Trejo, 1986; Trejo et al., 1987). A blue radar sweep and blue range indicators detected incoming hostile missiles and tracked their position with orange blips. Missiles appeared unpredictably from 360 degrees with at least 4 degrees between them. Each incoming missile appearing on the screen was assigned a random two-digit tracking number. Incoming missiles traveled toward the subject's ship which was at the center of the display, and was represented by a "+". The missiles traveled at three different speeds--slow, medium, and fast--with each missile speed constituting one-third of the missile total. The speed of any given missile was denoted by the spacing between its tracking blips; the faster the missile, the farther the blips were spaced apart.

The outer most circle represented the range of the ship's radar. The inner circle represented the maximum range of the ship's weapons. The subject's task was to avoid getting hit by the incoming missiles. A subject launched his missiles by moving a cursor over the desired tracking number with a mouse (an electronic hand-held cursor control device), and then pressing a button.

The ship's missiles were shown as green blips which traveled straight toward the incoming missiles at the speed of the fastest incoming missiles. If a subject launched too early on an incoming missile, his missile reached its maximum range before the incoming missile arrived and it "splashed." There was no penalty for a splash. On the other hand, if a subject failed to launch on an incoming missile, the incoming missile hit the ship. This was denoted by marking the track with an orange line. A "hit" resulted in a 12-point penalty. Only one missile could be launched on an incoming target at a time. If a subject attempted to launch on a target that already had a missile "in flight," a 2-point penalty resulted. A subject gained points based on a calculation of the average kill range for the incoming missiles. When a subject killed an incoming missile, its track was marked with a green line. The level of difficulty was controlled by varying the frequency of incoming missiles. The speeds of the individual missiles (fast, medium, and slow) did not vary with task difficulty.

There were some operational differences between the versions of AIRDEF used at NASNI and NCO-CP. A brief description of these differences follows.

Launching	In the NASNI experiment, subjects launched their ship's weapons by one of two methods. They either instructed the experimenter verbally as to the track number they wanted to fire on or they typed the track number on a keyboard themselves. In the NCO-CP experiment, subjects launched their ship's weapons by moving a cursor over the desired tracking number with a mouse (an electronic hand-held cursor control device), and then depressing a button.
Hits & Kills	In the NASNI experiment subjects were given feedback about an incoming missile hitting their ship by the appearance of two X's that crossed out the tracking number for that missile. If subjects killed an incoming missile, a solid orange line marked the trace left by that missile's blips. In the NCO-CP experiment, subjects received additional feedback about kills by the appearance of a solid green line marking that missile's blips.
Baseline	At NASNI the baseline condition was completely passive with regard to what the subjects were required to do. In this baseline condition no AIRDEF simulation activity occurred, and subjects simply fixated on a blank screen which intermittently presented probe stimuli. In the baseline condition at NCO-CP, the radar range indicators were displayed continuously. Subjects moved the launch cursor over randomly appearing blocks where tracking numbers usually appeared. Subjects pressed the mouse button as if to launch a weapon on these blocks, but no targets or weapons ever appeared on the screen. This baseline simulated the psychomotor activity that was required during actual performance of the task, but it required very little cognitive processing (speed estimation, memory, evaluation of tradeoffs, etc.) on the subjects' part, as did the active AIRDEF task.

Procedure

Subjects were fitted with an electrode helmet (Electro-Cap International[1]). Electrode impedances were usually within 1-2 Kohms and did not exceed 5 Kohms. Subjects were then seated approximately 45 cm away from the computer monitor in a non-reclining chair. At this point the subject was given instructions for AIRDEF from the experimenter.

At NASNI, subjects were told that a visual stimulus for producing EPs would appear intermittently during the baseline and active engagements of the AIRDEF task. They were also told that the probe had no relevance to the task and that it should not be attended to or counted.

Each experimental run consisted of three conditions; one passive low load (baseline) condition first, in which 10 visual probe stimuli were presented, but no simulation activity occurred. Next, each subject performed the AIRDEF simulation at a medium workload level with a target frequency of 4.5 targets per minute. Finally, each subject performed the simulation at a high workload level with a target frequency of 9 targets per minute. Each active engagement lasted for 4 minutes. During these engagements 20 probe stimuli were presented.

At NCO-CP, both visual and auditory oddball stimuli were presented. (The oddball task consists of presenting stimuli with different but complementary probabilities. The stimulus presented most often is referred to as the "frequent" stimulus and the stimulus presented least often is referred to as the "rare" stimulus.) The experimenter told the subjects that these stimuli would occur while they were performing the simulation and that they would be instructed by the computer as to which stimulus they should attend to (either visual or auditory rare). Subjects were

[1]Identification of the equipment is for documentation only and does not imply endorsement.

told that they would be called upon to report this number at the end of each engagement. Subjects were instructed to perform both the AIRDEF simulation (primary) and the counting (secondary) task to the best of their ability. When subjects attended to one modality (eg., audition), stimuli in the other modality (i.e., vision) were to be ignored. Thus, each engagement contained both secondary-task relevant probes and irrelevant probes as stimuli.

Each experimental run consisted of six conditions with three levels of difficulty. The three levels of difficulty consisted of a low, medium and high load. Six conditions were involved because subjects had to perform each difficulty level while attending to either a "dim" visual flash or a "low pitch" auditory tone. The low load condition was the baseline condition described earlier. The medium load condition had a target frequency of 3 missiles per minute. The high workload condition had a target frequency of 9 missiles per minute. Each engagement lasted 6 minutes. The order of difficulty presentation was counterbalanced. For more details, see Blankenship et al. (in press).

Subjects were then instructed to relax their jaw and face muscles as much as possible to minimize muscle artifacts while recording; the experimenter also asked the subjects to minimize eyeblinks.

Stimulus Presentation

NASNI

A visual irrelevant probe stimulus (non-oddball) was presented on the addressable area of the computer monitor (a rectangular flash covering the entire 13-in. computer monitor) used for the AIRDEF display. The visual stimulus was a background flash which did not interfere with the presentation of the task on the screen. When viewed from the average distance of 45 cm, the stimulus subtended 37° visual angle vertically and 49° horizontally. The stimulus duration was 16 ms, and its luminance was 10.3 nits. The mean inter-stimulus-interval (ISI) was 11 seconds with a range from 6 to 16 seconds.

NCO-CP

Visual and auditory oddball stimuli were presented during each level of AIRDEF difficulty with a .75 probability of a frequent stimulus occurring and a .25 probability of a rare stimulus occurring. This worked out to a total of 45 frequent and 15 rare stimuli for both the visual (60) and auditory (60) modalities (120 total). There was no overlap in stimulus presentation in either modality. On this basis, the mean ISI was 2 seconds (based on a 4-minute run time with 120 total stimuli for each condition). Since the topic of this report is only concerned with the closest replication possible, only the visual oddball stimuli presented in the attend auditory condition from NCO-CP will be considered beyond this point.

The visual oddball stimuli (45 frequent and 15 rare for each difficulty level) were presented on the computer monitor used for the AIRDEF display. These stimuli were identical to those used at NASNI except in luminance. The frequent stimulus luminance was 21.6 nits and the rare stimulus luminance was 5.1 nits.

Data Acquisition

The recording procedures in both experiments were similar. However, differences existed in the sampling frequency, reference electrode placement, and the number of epochs collected. The recording equipment consisted of an array of tin electrodes embedded in a nylon cap (Electro-Cap International), a set of amplifiers (Grass Model 12A5, Neurodata Acquisition System), and a computer (Masscomp Model 5500) programmed to digitize and record neuroelectric signals. The ERPs were bandpass filtered (analog, 3 dB corner frequencies at 0.1 and 100 Hz), amplified (20,000 gain), sampled at 256 Hz (NASNI) and 128 Hz (NCO-CP) and digitized. All ERP data were acquired for a period of one second and stored as single epochs. Recordings were made from eight sites: F3 and F4 (frontals), T3 and T4 (temporals), P3 and P4 (parietals), and O1 and O2 (occipitals) according to the International 10-20 System (Jasper, 1958). Recordings were also made from site Fp2 (right frontal pole) to monitor eye blinks and large eye movements. Fp2 is located just above the right eye and is capable of monitoring both horizontal and vertical eye movements. The electrical reference was Cz at NASNI and nose at NCO-CP. In both studies, the subject ground was at a point on the midline, approximately three cm anterior to the frontal (F3, F4) sites. All recorded sites were monitored on line with an eight-channel oscilloscope (Tektronix, Model 5103N).

Digital Filtering

After the data had been collected and stored, they were detrended (mean and linear slope removed) and digitally filtered (windowed finite impulse response filter, 127 coefficients, corner frequencies at 0.1 and 25 Hz, Hamming window).

Artifact Rejection

In order to remove eye blink and/or eye movement artifacts, three criteria were applied to the ERPs recorded at site Fp2. Any ERP with a transient signal exceeding a baseline-to-peak amplitude of 50 μV in the post-stimulus time period of 50 to 457 ms was rejected. These ERPs were then rejected at all recording sites. Rejection was done on the 50 to 457-ms post-stimulus window because this was the window to be analyzed.

An additional artifact rejection procedure was done at NCO-CP prior to the off-line rejection procedure outlined above. A storage oscilloscope was used for on-line rejection of bad epochs (Tektronix, Model 336). The criterion for epoch rejection was based on the ERPs recorded at site Fp2. ERPs that showed a transient signal with a baseline-to-peak amplitude greater than 50 μV during 500 ms of the post-stimulus period were rejected. If an ERP at Fp2 was rejected, then the corresponding ERPs at every recording site were rejected. Rejected ERPs were repeated to ensure the collection of the desired number of epochs for each stimulus type.

Selection of ERPs for Analysis

After artifact rejection, the number of artifact-free epochs varied between subjects. Three criteria were then applied to select artifact-free epochs for analysis. First, the number of epochs had to be equal across the three difficulty levels of AIRDEF (low load, medium load, and high load). Second, the number of epochs had to be equal across all subjects. Third, the number of epochs had to be one that would retain a majority of the subject sample for analysis. Based on those criteria, 30 subjects from NASNI had at least six artifact-free ERPs in the three difficulty levels of AIRDEF, and 65 subjects from NCO-CP had at least nine artifact-free ERPs in the three difficulty levels of AIRDEF.

Because workload (target frequency) in AIRDEF rises at the beginning of the engagement and falls near the end, ERPs recorded near the beginning or the end may not show workload-related effects. ERPs near the beginning and near the end of an AIRDEF trial were excluded from analysis whenever more than either six or nine artifact-free ERPs were available. This procedure resulted in retaining six ERPs that were generally from the middle of an AIRDEF engagement where workload is nearly constant.

Electrode Site Derivations

Because a different reference electrode placement was used in the recording array at NASNI and NCO-CP, a direct comparison could not be performed on the physiological data. This results from the possible contribution of activity near the reference to the ERP. In an attempt to make the data as compatible as possible, some electrode site derivations were computed from the data from each study to remove the effects of the reference electrode, according to the following formula,

$$E(Site_i - Site_j) = E(Site_i - Ref) - E(Site_j - Ref),$$

where E is the voltage at an electrode site, and Ref is the reference electrode. Based on this general derivation, any two sites that are subtracted from one another will yield a reference-independent derivation. Table 1 shows the site derivations that were computed on both the NASNI and NCO-CP electrode sites. As can be seen, six of the derivations concern ipsilateral sites and two of the derivations concern contralateral sites. These derivations were then submitted to the following signal processing procedures.

Signal Processing

From the number of ERPs selected for each subject, a signal-to-noise ratio waveform (SNWAVE) was computed (representing the 50 to 457-ms post-stimulus window). The SNWAVE was computed by dividing the signal averaged waveform by its standard deviation on a point-by-point basis. See Trejo et al. (1987) for a detailed discussion of this measure.

The SNWAVEs for each subject were then divided into eight equal latency windows, each 50.78 ms wide (13 sample points at 256 Hz). Because the sampling frequency at NCO-CP was 128 Hz and at NASNI it was 256 Hz, a linear interpolation was done on the NCO-CP data to best match the latency windows. Table 2 lists the beginning,

Table 1

Electrode Site Derivations	
Number	Derivation
1	F3 - P3
2	F3 - T3
3	F3 - P4
4	O1 - P3
5	F4 - P3
6	F4 - T4
7	F4 - P4
8	O2 - P4

end, and center latencies for these eight windows. By using a standard set of time windows, a dependent amplitude measure could be computed for each individual for each time window.

For each of these windows, an unbiased root-mean-square (RMS) integrated amplitude measure was computed in µV units. These RMS amplitude values were used as the dependent measure in statistical analyses. This particular amplitude measure has been used and discussed elsewhere (Callaway, 1975; Callaway, Halliday, & Herning, 1983; Lewis & Sorenson, 1987; Trejo et al., 1987).

RESULTS

A three-way repeated measures analysis of variance (ANOVA) was performed separately on the NASNI and NCO-CP data. The independent factors were workload, site derivation, and time window.

(1) The workload factor had three levels assessed by target frequency on the AIRDEF simulation. The first level was the baseline condition, while the second two levels involved incoming targets. At NASNI these two engaged levels consisted of 4.5 and 9 incoming targets per minute, while at NCO-CP the two engaged levels consisted of 3 and 9 targets per minute.

(2) The site factor had eight levels comprised of the derivations outlined in Table 1.

(3) The window factor was comprised of the eight latency windows. These windows were described in Table 2.

The dependent variable was the RMS amplitude for the discrete windows. Results from these analyses are summarized in Table 3. In order to minimize the opportunity for taking the advantage of chance in this validation, we are only considering those effects which were found to be significant in the first NASNI report by Trejo et al. (1987). In that report, an ANOVA was conducted on non-derived sites as in the ANOVA reported here for the derived sites. As outlined above, the methods for ERP selection and analysis were the same. Briefly, it was found

Table 2

Time Windows for Analyses of ERP Waveforms			
Window	Window Latency (ms)		
Number	Beginning	End	Midpoint
1	50.78	101.56	76.17
2	101.56	152.34	126.95
3	152.34	203.13	177.74
4	203.13	253.91	228.52
5	253.91	304.69	279.30
6	304.69	355.47	330.08
7	355.47	406.25	380.86
8	406.25	457.03	431.64

that the main effect fo. workload and the interaction of workload by site by time were significant in the analysis. Based on those findings we will only consider those effects in the ANOVA reported here.

As can been seen in part A of Table 3, both effects of interest are significant. The workload main effect was $F(2, 58) = 5.54$, $p < .0128$, and the three-way interaction was $F(98, 2842) = 1.38$, $p < .0083$. Similar results were found for the NCO-CP data reported in part B of Table 3. The workload main effect was $F(2, 128) = 3.74$, $p < .0264$, and the three-way interaction effect was $F(98, 6272) = 1.78$, $p < .0078$. All of the reported p values have been corrected for violations of sphericity. These corrected values represent the Greenhouse-Geisser correction (BMDP Statistical Software, 1985).

Simple effects analyses were carried out on the three-way interaction means (workload by site derivation by time). These simple effects are reported in Table 4. Again, the reported p values have been corrected for violations of sphericity (Greenhouse-Geisser correction, BMDP Statistical Software, 1985). Across the two studies results replicated for three time windows (4, 6 and 7) at four derivations (F3-P3, F3-P4, O1-P3, and O2-P4). These results are shown in bold face.

DISCUSSION

The objective of this report was to determine if ERP amplitude decreases in relation to cognitive workload in one study could be replicated in a second study. The results support this objective at both global and specific levels. The general findings suggest that particular electrode site and time (window latency) combinations do, in fact, change systematically with increasing task demands. The findings suggest that the effect is robust, since the replication was conducted under different experimental manipulations.

Table 3

A. ANOVA Summary Table for NASNI Derivations

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	14.73	2	7.37	5.54	.0128
W × Subjects (Sb)	77.14	58	1.33		
Site Derivation (Sd)	84.94	7	12.13	90.85	.0000
S × Sb	27.11	203	0.13		
Time (T)	8.84	7	1.26	4.17	.0019
T × Sb	61.46	203	0.30		
W × Sd × T	6.31	98	0.06	1.38	.0083
W × Sd × T × Sb	132.18	2842	0.05		

B. ANOVA Summary Table for NCO-CP Derivations

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	2.46	2	1.23	3.74	.0264
W × Subjects (Sb)	42.05	128	0.33		
Site Derivation (Sd)	203.80	7	29.11	149.95	.0000
S × Sb	86.98	448	0.19		
Time (T)	184.11	7	26.30	78.19	.0000
T × Sb	150.70	448	0.34		
W × Sd × T	7.10	98	0.72	1.78	.0078
W × Sd × T × Sb	254.90	6272	0.04		

Table 4

A. Simple Effects Analyses for NASNI Derivations

Site Derivation	Window Number	RMS Signal-to-Noise Mean Baseline	Level I	Level II	*F*	*p*
F3-P3	4	.68	.45	.49	3.16	.058
	6	**.65**	**.45**	**.51**	**3.68**	**.039**
F3-T3	2	.43	.28	.34	4.09	.022
F3-P4	**4**	**.75**	**.53**	**.52**	**3.10**	**.059**
	6	.74	.48	.40	10.13	.001
O1-P3	6	.41	.25	.19	6.24	.011
	7	**.33**	**.26**	**.19**	**5.08**	**.016**
F4-P3	4	.74	.47	.51	3.96	.032
	6	.64	.40	.49	5.78	.008
F4-T4	2	.52	.36	.37	4.17	.027
	4	.53	.31	.35	6.74	.005
F4-P4	2	.87	.58	.66	4.89	.015
	4	.73	.46	.48	4.46	.021
O2-P4	3	.42	.29	.28	3.42	.056
	6	.36	.21	.22	4.59	.023
	7	**.34**	**.24**	**.20**	**5.36**	**.013**

B. *Simple Effects* Analyses for NCO-CP Derivations

Site Derivation	Window Number	RMS Signal-to-Noise Mean Baseline	Level I	Level II	*F*	*p*
F3-P3	1	.54	.43	.49	6.10	.003
	5	.39	.32	.34	5.21	.008
	6	**.47**	**.40**	**.40**	**4.25**	**.016**
F3-P4	**4**	**.44**	**.40**	**.35**	**7.94**	**.001**
	5	.38	.34	.33	3.13	.049
O1-P3	2	.92	.82	.93	4.06	.020
	4	.49	.62	.61	6.97	.002
	5	.56	.59	.51	3.30	.043
	7	**.44**	**.36**	**.40**	**6.69**	**.003**
F4-P3	7	.44	.39	.37	3.76	.027
O2-P4	4	.53	.64	.66	4.90	.009
	7	**.47**	**.38**	**.39**	**5.04**	**.008**

At the global level of analysis, results from the ANOVAs indicated that ERPs changed significantly as a function of workload in both the single-task NASNI experiment ($F(2, 58) = 5.54$, $p < .0128$) and in the dual-task NCO-CP experiment ($F(2, 128) = 3.74$, $p < .0264$). This change occurred between the baseline and engaged conditions in both studies. Similar results have been reported by others who have used ERPs in assessing cognitive workload (Isreal, Chesney, Wickens, & Donchin, 1980; Trejo et al., 1987). In both of the studies compared in this report there was a significant three-way interaction between workload, site derivation, and time (see Table 3). The simple effects of this interaction indicated that specific 50-ms wide windows in particular site derivations could be

reliably replicated across studies. The derivations included F3-P3, F3-P4, O1-P3, and O2-P4. All these derivations include sites reported by Trejo et al., (1987) as being significantly related to workload changes in the AIRDEF task. Two of the three time windows that were replicated were also significant in the Trejo et al. study.

The three replicated windows (4, 6, and 7) spanned a post-stimulus period from 203-406 ms, indicating that the observed results were more reliable with later cognitive components. Specifically, the fourth window was probably representative of a P2 component, while windows six and seven were representative of the P300 component. The mean decrease in the signal-to-noise ratio from the baseline condition to the average of the two engaged conditions at NASNI (for the replicated simple effects) was 30.8 percent, with a range from 26 to 35 percent. This same mean decrease at NCO-CP was 15.5 percent, with a range from 14 to 18 percent. The differences in the amount that the ratio decreased, as well as the mean differences, could be due to the different experimental conditions under which the data were collected. The important finding is that the changes were in the same direction.

The signal-to-noise ratio for the occipital-parietal derivations (O1-P3, O2-P4) increased significantly as a function of workload in the NCO-CP study. These increases contradict the predictions of a simple, undifferentiated capacity model of cognitive resource allocation. They also contradict one alternative explanation of the decreases that occurred at other site-window combinations; that is, that general excitability of the nervous system decreases when subjects are actively engaged in complex tasks as compared to passive or simple tasks. One possible explanation for the occipital-parietal window 4 signal-to-noise ratio increases arises from the increased level of attention to visual details in the active AIRDEF engagements as compared to baseline. Attention effects have been related to increased amplitudes of the peak-to-peak amplitude of the N1-P2 complex (Näätänen, 1982), which may overlap temporarily with window 4.

None of the earlier windows (less than 200 ms) were replicated across studies. Within each study, however, there were significant simple effects within earlier time windows. These earlier windows have been mostly associated with sensory processing of stimuli. If these early windows index sensory processing, which depends strongly on stimulus features, then the lack of replication is understandable. The stimuli between the two studies differed in their luminance, context, and frequency of occurrence. The striking result, then, is the finding that in both studies ERPs are reliable indices of cognitive workload.

The results that were replicated are congruent with other ERP studies concerned with the assessment of cognitive workload (Blankenship et al., in press; Gopher & Donchin, 1986; Kramer, Wickens, & Donchin, 1986). In most of the dual-tasking studies that have been conducted using cognitive psychophysiological techniques to assess mental workload, the primary task included complex behavioral performance, while the secondary task consisted of an oddball paradigm (Defayolle, Dinard, & Gentil, 1971; Isreal et al., 1980; Karis, Coles, & Donchin, 1984; Kramer, Wickens, & Donchin, 1986). Few psychophysiological studies have used irrelevant probe stimuli in attempts to study cognitive processes (Papanicolaou & Johnstone, 1984; Trejo et al., 1987). The replication reported here adds weight to the irrelevant probe method as a means of cognitive assessment since the results from the NASNI experiment were similar to those from the NCO-CP dual-tasking (oddball) experiment. Future experiments should explore the irrelevant probe paradigm in greater detail. If, in fact, this method does allow for the assessment of cognitive processing, it will be an improvement over the dual-task paradigm because it interferes less with the primary task than does the oddball technique.

REFERENCES

Blankenship, M., Trejo, L., & Lewis, G. (in press). *Brain activity during tactical decision-making: IV. Event-related potentials as indices of selective attention and cognitive workload* (NPRDC Tech. Note 88-xx). San Diego: Navy Personnel Research and Development Center.

BMDP Statistical Software. (1985). *P2V--Analysis of variance and covariance including repeated measures.* Berkeley: University of California Press.

Callaway, E. (1975). *Brain electrical potentials and individual psychological differences.* New York: Grune & Stratton.

Callaway, E., Halliday, R., & Herning, R. I. (1983). A comparison of methods for measuring event-related potentials. *Electroencephalography and Clinical Neurophysiology, 55,* 227-232.

Defayolle, M., Dinard, J. P., & Gentil, M. T. (1971). Averaged evoked potentials in relation to attitude, mental load and intelligence. In W. T. Singleton, J. G. Fox, & D. Whitfield (Eds.), *Measurement of man at work.* New York: Van Nostrand Reinhold.

Fishman, J. A. (1958). Unsolved criterion problems in the selection of college students. *Harvard Educational Review, 28* (4), 340-349.

Ghiselli, E. E. (1966). *The validity of occupational aptitude tests.* New York: John Wiley.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26,* 461-477.

Gopher, D., & Donchin, E. (1986). Workload--An examination of the concept. In K. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II.* New York: John Wiley.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17* (3), 259-273.

Jasper, H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 211-224.

Karis, D., Coles, M. G. H., & Donchin, E. (1984). *The use of event related brain potentials for prediction of performance quality in pilots of advanced tactical jet aircraft* (Tech. Rep. CPL 84-2/F33615-82-K-0611). Champaign: CPL, University of Illinois.

Kelly, R. T., Greitzer, F. L., & Hershman, R. L. (July 1981). *Air Defense: A computer game for research in human performance* (NPRDC Tech. Rep. 81-15). San Diego: Navy Personnel Research and Development Center.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1986). Processing of stimulus properties: Evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance, 11* (4), 393-408.

Lewis, G. W. (1983). *Bioelectric predictors of personnel performance: A review of relevant research at the Navy Personnel Research and Development Center* (NPRDC Tech. Rep. 84-3). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W. (1984). Temporal stability of multichannel, multimodal ERP recordings. *International Journal of Neuroscience, 25*, 131-144.

Lewis, G. W., & Rimland, B. (1979). *Hemispheric asymmetry as related to pilot and radar intercept officer performance*. (NPRDC Tech. Rep. 79-13). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W., & Sorenson, R. C. (1987). *Evoked brain activity and personnel performance* (NPRDC Tech. Rep. 88-1). San Diego: Navy Personnel Research and Development Center.

Näätänen, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. *Psychological Bulletin, 92*, 605-640.

Papanicolaou, A. C., & Johnstone, J. (1984). Probe evoked potentials: Theory, method and applications. *International Journal of Neuroscience, 24*, 107-131.

Trejo, L. J. (April 1986). *Brain activity during tactical decision-making: I. Hypotheses and experimental design* (HFOSL Tech. Note 71-86-6). San Diego: Navy Personnel Research and Development Center.

Trejo, L. J., Lewis, G. W., & Blankenship, M. H. (1987). *Brain activity during tactical decision-making: II. Probe-evoked potentials and workload* (NPRDC Tech. Note 88-12). San Diego: Navy Personnel Research and Development Center.

DISTRIBUTION LIST

Office of the Chief of Naval Operations (OP-933D4)
Deputy Under Secretary of Defense for Research and Engineering
Assistant for Manpower Personnel and Training Research and Development (OP-01B2)
Director, Command Surface Warfare Manpower and Training Requirements (OP-39)
Naval Training Systems Center, Technical Library (5)
Head, Cognitive and Neurosciences (OCNR-1142)
Life Sciences Directorate (OCNR-114)
Perceptual Science (OCNR-1142PS)
Cognitive and Decision Science (OCNR-1142CS)
Technology Area Manager, Office of Naval Technology (Code 222), (Code 223)
Commanding Officer, Naval Aerospace Medical Research Laboratory, Pensacola, FL (Code 031)
Commanding Officer, Naval Health Sciences Education and Training Command, Bethesda, MD
Naval Medical Research and Development Command (Code 40)
Naval Medical Command (02D)
Naval Biodynamics Laboratory
Commanding Officer, Naval Hospital, Portsmouth, VA (Medical Library)
Army Research Institute (PERI-SB)
Headquarters, Marine Corps (Code MA)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base, TX
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Program Manager, Manpower Research and Advisory Service, Smithsonian Institute
Center for Naval Analyses, Acquisitions Unit
Defense Technical Information Center (DTIC) (2)